EXHIBIT A

Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 878–883, January 1996
Biophysics

Three-dimensional structure of recombinant human osteogenic protein 1: Structural paradigm for the transforming growth factor β superfamily

(x-ray crystallography/protein structure/bone morphogenetic protein/cystine knot/receptor binding model)

DIANA L. GRIFFITH*†, PETER C. KECK‡, T. KUBER SAMPATH‡, DAVID C. RUEGER‡, AND WILLIAM D. CARLSON§

*Rosenstiel Basic Medical Research Center, Brandeis University, Waltham, MA 02254; ‡Creative BioMolecules Inc., Hopkinton, MA 01748; and §Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115

Communicated by Gregory A. Petsko, Brandeis University, Waltham, MA, October 3, 1995

ABSTRACT **We report the three-dimensional structure of osteogenic protein 1 (OP-1, also known as bone morphogenetic protein 7) to 2.8-Å resolution. OP-1 is a member of the transforming growth factor β (TGF-β) superfamily of proteins and is able to induce new bone formation *in vivo*. Members of this superfamily share sequence similarity in their C-terminal regions and are implicated in embryonic development and adult tissue repair. Our crystal structure makes possible the structural comparison between two members of the TGF-β superfamily. We find that although there is limited sequence identity between OP-1 and TGF-β2, they share a common polypeptide fold. These results establish a basis for proposing the OP-1/TGF-β2 fold as the primary structural motif for the TGF-β superfamily as a whole. Detailed comparison of the OP-1 and TGF-β2 structures has revealed striking differences that provide insights into how these growth factors interact with their receptors.**

Osteogenic protein 1 [OP-1, also called bone morphogenetic protein 7(BMP-7)] was originally isolated from bone based on its ability to induce new bone formation *in vivo* (1). Preclinical studies in nonhuman primate models have demonstrated that OP-1 is effective in restoring large segmental bone defects (2). Evidence that OP-1 is synthesized in the kidney (3) and is present in the circulation suggests that OP-1 has therapeutic potential for the treatment and management of osteoporosis and related metabolic bone diseases. In addition, tissue localization and other preclinical studies suggest that OP-1 has a role in the repair and regeneration of urogenital (4), neuronal (5), and cardiovascular (6) tissues. Knowledge of the three-dimensional structure of OP-1 is essential for understanding its mode of action and for providing a basis for the development of small molecule therapeutics.

OP-1 and transforming growth factor β2 (TGF-β2) are 2 of >30 homologous proteins in the TGF-β superfamily. Members of this superfamily have diverse biological activities and play critical roles in the migration, proliferation, and differentiation of mesenchymal cells during embryogenesis and in the repair and regeneration of tissues during postfetal life (7). These proteins are synthesized as large precursor proteins that undergo proteolytic processing at RXXR sites to yield mature active dimers of disulfide-linked monomers. Each monomer contains a conserved C-terminal 7-Cys domain with 20–92% sequence identity among superfamily members (7).

In this paper we report the three-dimensional structure of mature OP-1 to 2.8-Å resolution.[¶] The overall OP-1 monomer fold is a Greek key motif with approximate dimensions 60 Å × 20 Å × 15 Å. In large part, the analogy to a left hand for TGF-β2 is true for OP-1 as well (8) (Figs. 1 and 2). The overall

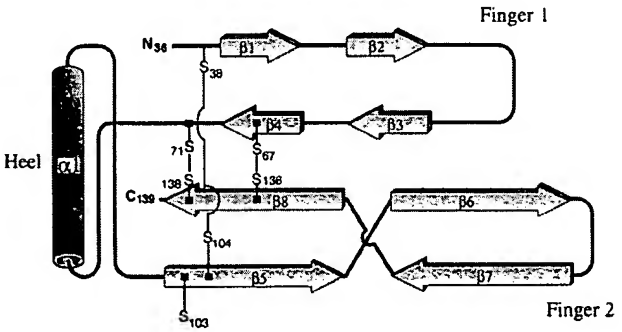


FIG. 1. Schematic drawing of the OP-1 monomer fold. The OP-1 cystine knot constitutes the core of the monomer and consists of three disulfide bonds; two, Cys-67–Cys-136 and Cys-71–Cys-138, form a ring through which the third, Cys-38–Cys-104, passes. The four strands of antiparallel β-sheet, which emanate from the knot, form two finger-like projections. An α-helix, located on the opposite end of the knot, lies perpendicular to the axis of the two fingers thereby forming the heel of the hand. The N terminus that corresponds to the thumb of the hand is unresolved in our electron density map. Unlike TGF-β2 (8, 9), this N-terminal region is not stabilized by a disulfide bond. The β-sheets are displayed as arrows and labeled from β1 through β8. The α-helix is displayed as a tube and labeled α1. Shown in solid thin lines are the intrasubunit disulfide bonds that make up the cystine knot. Starting from Gln-36, the residues involved in regular secondary structure are β1 (Lys-39 to His-41), β2 (Tyr-44 to Ser-46), β3 (Glu-60 to Ala-63), and β4 (Tyr-65 to Glu-70) in finger 1; β5 (Cys-103 to Asn-110), β6 (Ile-112 to Asp-118), β7 (Asn-122 to Tyr-128), and β8 (Val-132 to His-139) in finger 2; and α1 (Thr-82 to Ile-94) in the heel.

architectural similarity between OP-1 and TGF-β2 (8, 9) has allowed us to construct a structure-based sequence alignment (Fig. 3) from which we compared these two structures to begin to understand the chemical and structural elements involved in determining the specificity of these proteins to their receptors.

METHODS

Structure Determination. Crystals of human recombinant mature OP-1 were grown by mixing equal volumes of purified protein (10, 11), at 10 mg/ml, with 8% saturated ammonium sulfate in 50 mM sodium acetate (pH 5.0) (12). The crystals

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: OP-1, osteogenic protein 1; BMP, bone morphogenetic protein; TGF-β, transforming growth factor β.

†To whom reprint requests should be addressed at: Leukocyte Biology and Inflammation Program, Renal Unit, Massachusetts General Hospital and Department of Medicine, Harvard Medical School, 149 13th Street, Charlestown, MA 02129.

¶The atomic coordinates and structure factors have been deposited in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973 (reference BMPI). This information is embargoed for 1 year (coordinates) from the date of publication.

FIG. 2. Stereoview of the OP-1 dimer, viewed along the crystallographic twofold axis from the N-terminal side. The single intermolecular disulfide bond linking Cys-103 of each monomer is shown in yellow and lies exactly on this twofold axis. The OP-1 monomer has an accessible nonpolar surface area of 4394 Å^2, and the dimer has an accessible nonpolar surface area of 6831 Å^2, resulting in a hidden area upon dimerization of 979 Å^2 per monomer. We similarly calculate values of 4448 Å^2 for the monomer and 6852 Å^2 for the dimer of the TGF-β2 structure 2TGI (8) and 4412 Å^2 for the monomer and 6779 Å^2 for the dimer of the TGF-β2 structure 1TFG (9). In both cases, the hidden area upon dimerization is 1022 Å^2 per monomer.

have the symmetry of space group $P3_221$ with unit-cell dimensions of $a = b = 99.46$ Å and $c = 42.09$ Å. One crystal was used to collect a complete native data set to 2.8-Å resolution at 4°C. Two heavy atom derivative data sets were collected at 4°C, one from a crystal soaked for 7 days in 0.3 mM uranyl nitrate and the other from a crystal soaked for 8 h in 0.5 mM sodium gold (III) tetrachloride (12) (Table 1). The native and derivative data sets were integrated and reduced with the R-AXIS-IIC software suite (13) and scaled together with the CCP4 program ANSC (14). Inspection of the Harker sections of the difference Patterson map revealed a single uranyl site. The position of the single gold site was determined by employing cross-Fourier techniques using the uranyl position as the phasing site. The heavy atom x,y,z parameters and occupancy were refined with the program TENEYCK (15). By using these two derivatives and their anomalous signals, an initial phase set was calculated to 4.0-Å resolution with a mean figure of merit of 0.72. The phases were improved and extended to 3.5-Å resolution by cycles of solvent flattening (16) and phase combination (17) using the CCP4 (14) crystallographic package. A completely interpretable 3.5-Å resolution electron density map permitted the unambiguous tracing of the polypeptide chain and identification of the amino acids from Gln-36 to His-139 by using the graphic program O (18). The model was refined with the program XPLOR (19) by using all reflections between 10-Å and 2.8-Å resolution for which $F_{obs} > 2.0\sigma(F_{obs})$. There were no water molecules included in the refinement. The rms deviation from ideality is 0.02 Å for bond lengths and 3.2° for bond angles. The current R factor is 22.8%.

Structural Alignment of Sequences and Calculation of Structure-Based Sequence Identity. The sequences of OP-1 and TGF-β2 were aligned in Fig. 3 according to the corre-

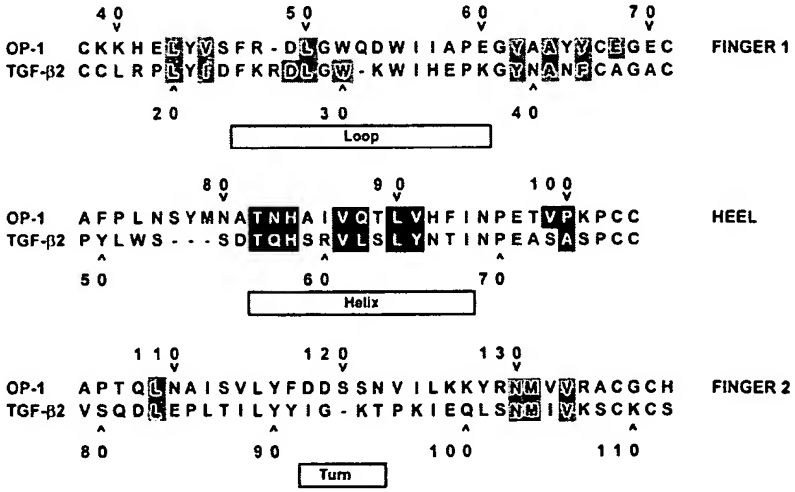


FIG. 3. Structure-based sequence alignment of the OP-1 and TGF-β2 7-Cys domains. Residues from the heel region involved in interchain contacts in the dimers of OP-1 and TGF-β2 are white type on black backgrounds, while the residues that they contact in the finger 1 region (upper row) and finger 2 region (lower row) of the other chain are black type on shaded backgrounds. In both OP-1 and TGF-β2, the same residue positions are involved in making interchain contacts.

Table 1. OP-1 structure determination parameters

Statistics	Native	Gold	Uranyl
Data collection			
Completeness to 2.8 Å, %	87	84	92
$+R_{merge}$, %	6.7	10.0	6.0
Phasing and refinement			
$*R_{iso}$, %		12.5	10.0
Heavy atom sites, no.		1	1
Isomorphous reflections, no.		3590	1863
Anomalous pairs collected, no.		2852	3697
Mean figure of merit		0.52	0.69
Resolution, Å	2.8		
Water molecules, no.	0		
R factor, %	22.8		
rms bond length, Å	0.02		
rms bond angle, degrees	3.2		

$+R_{merge} = \Sigma (|I - \{I\}|)/\Sigma I$, where I is the intensity measurement for symmetry-related reflections and $\{I\}$ is the mean intensity. $*R_{iso} = \Sigma h|F_{PH} - F_P|/\Sigma hF_P$, where F_P and F_{PH} are, respectively, the native and derivative structure factor amplitudes.

sponding regions of local structural identity in the OP-1 and TGF-β2 structures. Alignment gaps were positioned in loop regions, which is where the local conformational homology of the α-carbon traces tends to be the lowest. The structure-based alignment of OP-1 and TGF-β2 was then used as a template for the alignment of the 7-Cys domain sequences of other TGF-β superfamily members. Alignment gaps were positioned in regions that are loops in both the OP-1 and TGF-β2 structures. Percent identity between pairs of sequences was calculated as the number of identical aligned sequence positions, excluding gaps, normalized to the geometric mean of the lengths of the sequences, and multiplied by 100.

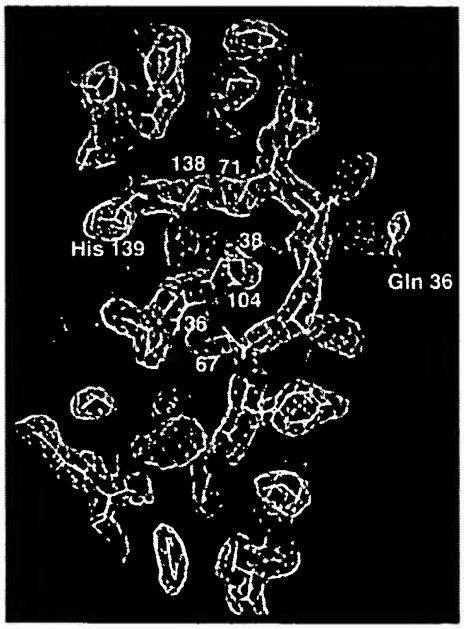


FIG. 4. Protein model of the cystine knot superimposed on the corresponding 2Fo − Fc electron density map. The peptide backbone for amino acid residues Cys-67 to Cys-71 runs vertically up the right of the figure, while the peptide backbone for residues Cys-136 to Cys-138 runs up the left side. The core disulfide, Cys-38–Cys-104, passes through the ring in the middle of the figure with Cys-104 on the near left side and Cys-38 on the far right side. Sulfur atoms are depicted in green with disulfide bonds absent. The central position on the long side of the knot is a Gly residue (Gly-69) that is conserved among cystine knot structures (8, 9, 25–27) and the TGF-β superfamily sequences. Gly is the only amino acid able to adopt the atypical conformation necessary to accommodate the packing of this structure (9). This figure was rendered by using the program O (18).

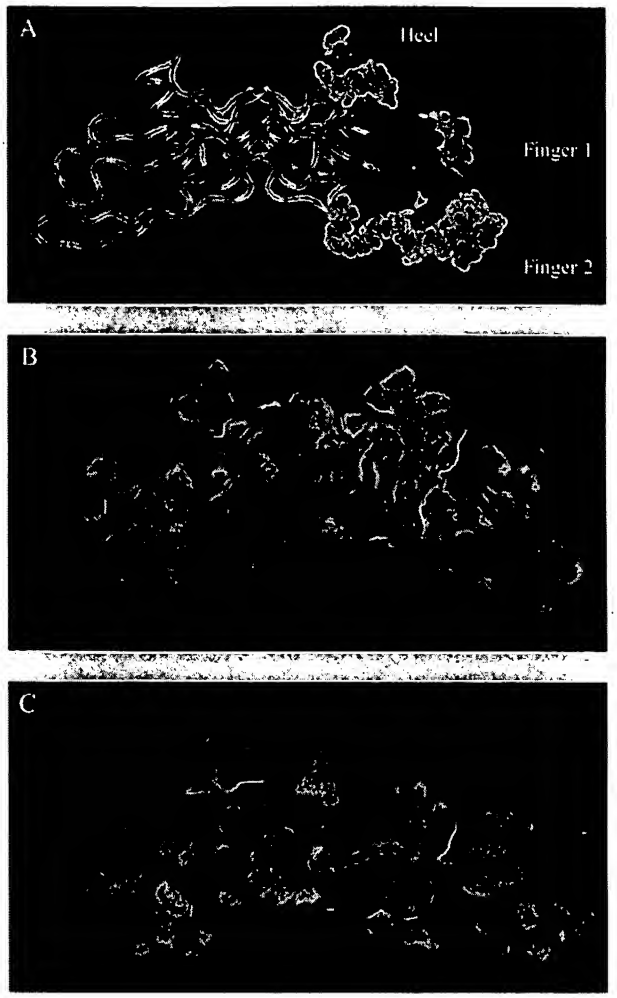


FIG. 5. (*A*) Ribbon rendition of the OP-1 dimer displaying the two monomer chains, one in green and the other in gold. Residues in the region of the proposed receptor binding ridge with solvent-exposed side chains are rendered as atomic spheres. The tips of fingers 1 and 2 of one chain and the loop at the C-terminal end of the helix of the other chain make up this ridge. Residues that are located at positions of variable sequence composition are white while the more conserved residues are red. (*B* and *C*) The solvent-accessible surfaces of the OP-1 (*B*) and TGF-β2 (*C*) dimers are colored based on their electrostatic potential. Surface regions having an electrostatic potential of $-3\ kT$ or less are red while surface regions of $+3\ kT$ or greater are blue. Neutral regions are green or gold to correspond to the backbone ribbons shown in *A*. This figure was rendered by using the program INSIGHT II (Biosym Technologies, San Diego).

Residue Contacts. In Fig. 3 residues were identified as contacting if the distance between the centers of at least one nonhydrogen atom from each side chain was less than the sum of their van der Waals radii plus 1.1 Å.

Angle and Dihedral Angle Between the Axes of the Helices in the Dimer. The helical axis was defined as the line equidistant from the α-carbons in the helical region. A sequence of four points is needed to define the dihedral angle between the axes of the helices in the dimer. The two inner points were chosen to lie on the helical axes adjacent to the α-carbon of residue His-84 in OP-1 or His-58 in TGF-β2, respectively. The two outer points were chosen to lie on their respective helical axes, but their location is arbitrary. To measure the angle between the helices, the first two points used to define the dihedral angle were translated to superimpose the inner points. The resulting three points define the angle.

Surface Area Measurements. The amount of nonpolar surface area lost during the formation of the dimer was calculated by using ACCESS (version 2.1) with a 1.4-Å probe (20). Nonpolar surface area is defined as the contribution to the total accessible surface from carbon and sulfur atoms. The surface area measurement algorithm in ACCESS slices the structure into 0.25-Å slabs perpendicular to the z axis. As a consequence, the results are sensitive to the orientation of a structure relative to the z axis (20). To minimize this effect, we evaluated three perpendicular and one intermediate orientations of each structure. The results of these calculations were combined by accepting, for each nonpolar atom, the largest accessible area measured among the four orientations. The values for TGF-β2 reported here were calculated by using coordinates from entry 2TGI (8) and entry 1TFG (9) obtained from the January 1994 release of the Protein Data Bank (21) at Brookhaven National Laboratory.

Electrostatic Calculations. The solution electrostatic potentials surrounding the OP-1 and TGF-β2 (1TFG) (9) dimers were calculated by using DELPHI (22, 23) (Biosym Technologies, San Diego). The calculations were performed by using a solvent dielectric constant of 80, a solvent radius of 1.4 Å, an ionic strength of 0.145 M, and an ionic radius of 2.0 Å. The interior of the protein was modeled by using a dielectric constant of 2.0. Formal charges were used and distributed as follows: atoms OD1 and OD2 of Asp were each charged −0.5, atoms OE1 and OE2 of Glu were each charged −0.5, atoms ND1 and NE2 of His were each charged 0.25, atom NZ of Lys was charged +1.0, atoms NH1 and NH2 of Arg were each charged +0.5, and atom OXT of the C-terminal carboxyl group was charged −1.0.

RESULTS

The Monomer Fold. The conserved 7-Cys domain of the TGF-β superfamily corresponds in OP-1 and TGF-β2 (8, 9) to a ring structure termed the cystine knot (24). Similar structures have been observed in several other growth factors that are not part of the TGF-β superfamily (25–27). The role of the cystine knot in these growth factors is to define the trifold topology of the monomer, which allows for structural variation to occur in the external loops (24). Fig. 4 shows a representative section of the OP-1 electron density map in the region of the cystine knot.

The finger 1 region of OP-1 is an antiparallel β-sheet containing a 13-residue Ω loop (Phe-47–Glu-60) (28) (Fig. 1). The structural alignment of the OP-1 and TGF-β2 sequences in Fig. 3 places two gaps in the Ω loop. The first gap represents a deletion in OP-1 that aligns with Arg-26 in the α2-helix of TGF-β2. This deletion results in a tighter non-α-helical turn in OP-1 as compared with TGF-β2. The second gap corresponds to the insertion of Gln-53 in OP-1, which has the result of directing both Gln-53 and Asp-54 side chains into the solvent. By comparison, in the corresponding region of TGF-β2, only Lys-31 is in contact with the solvent. These differences in the conformation of the Ω loop also result in the conserved proline (Pro-59) adopting a trans conformation in OP-1 rather than cis, as in TGF-β2 (9). The conformation of the Ω loop orients six nonpolar residues so they can contribute to a solvent-inaccessible interface with finger 2. Of these six, four are aromatic (Phe-47, Trp-55, Tyr-62, and Tyr-65), and two are aliphatic (Ile-56 and Ile-57). In all, the conformation of the Ω loop backbone places five polar residues (Arg-48, Asp-49, Gln-53, Asp-54, and Glu-60) in contact with solvent. The net surface charge in this region is −2, whereas it is +2 for TGF-β2 (Fig. 5).

The only α-helix in the monomer is located between the third and fifth Cys residues (Cys-71 and Cys-104) (Fig. 1). This helix extends for 3.5 turns from residues Thr-82 to Ile-94, is amphipathic, and contains a number of hydrophobic residues that in the dimer make contact with residues from finger 1 and finger 2 of the other monomer (Fig. 3). Several hydrophilic residues (Thr-82, His-84, and Gln-88) form one wall of an internal solvent pocket near the twofold axis of the dimer, while others (Asn-83, His-92, and Asn-95) are in contact with the external solvent. The conformation of the loop leading from the C-terminal end of the helix back to the cystine knot is similar in OP-1 and TGF-β2. By comparison, the loop located at the N-terminal end of the helix is 3 residues longer in OP-1, resulting in a different fold than in TGF-β2. In OP-1 this loop has an N-linked sugar moiety attached (Asn-80) (R. Tucker, personal communication); there is no corresponding glycosylation site in TGF-β2. Further, this loop is uncharged in OP-1, whereas it is negatively charged in TGF-β2.

Finger 2 is the second antiparallel β-sheet in OP-1 (Fig. 1). The polypeptide chain reverses direction between segments β6 and β7 through a 3:5 turn (29) beginning at residue Asp-118 and ending at residue Asn-122. In contrast, TGF-β2 has one less residue in this loop and adopts a 2:2 turn (29). Residues Arg-129 to Val-132, located between segments β7 and β8, form a peptide bridge that crosses over the C-terminal end of strand β5 and produces a 180° twist in the finger 2 antiparallel β-structure. A similar structure is observed in other cystine knot growth factors; however, the peptide bridge length varies (30). Within the monomer, finger 2 makes intrachain contacts with finger 1 by contributing aromatic residues Tyr-116, Phe-117, and Tyr-128 and aliphatic residues Val-114, Leu-115, Val-123, Met-131, and Val-133 to a solvent-inaccessible interface. OP-1 and TGF-β2 differ by three charges in the region of the finger 2 turn; OP-1 has two negative charges while TGF-β2 has one positive charge. In the region between the turn and the peptide bridge, OP-1 has a net charge of +3 while TGF-β2 is neutral (Fig. 5).

Dimer Fold. The OP-1 dimer is formed by the association of the heel region α-helix of one monomer with the finger regions of the other monomer and is stabilized by a single interchain disulfide bond involving Cys-103 (Figs. 1 and 2). Formation of this disulfide bond brings the cystine knot regions close together with the monomers oriented such that the molecular twofold axis passes between and parallel to the rings of the knots and perpendicular to the axes of the fingers, as in TGF-β2 (8, 9). In addition to the similarity of subunit orientation, formation of the OP-1 dimer hides nearly equal areas of nonpolar surface (Fig. 2).

A major difference between the OP-1 and TGF-β2 dimers is the relative orientation of the helices in the heel region. The angle between the axes of the helices in the heel region of OP-1 is 43°, which is 10° larger than that measured for TGF-β2. The measured dihedral angle between the helices is −20° for OP-1, which is 14° more negative than for TGF-β2. Despite these differences in helical orientation, the same helix and finger residue positions are involved in making interchain contacts, as evidenced by the shaded residues in Fig. 3.

DISCUSSION

OP-1 as a Structural Paradigm for the TGF-β Superfamily. The OP-1 structure presented here confirms the proposal that the TGF-β2 fold is prototypical for the conserved C-terminal region of the TGF-β superfamily (8, 9). We have aligned the sequences of the C-terminal 7-Cys domain of a number of TGF-β superfamily proteins by using the structurally based sequence alignment of OP-1 and TGF-β2 as a template (Fig. 3). The sequence of the finger 1 Ω loop in OP-1 has higher homology with other TGF-β superfamily members than does the corresponding sequence in TGF-β2. Specifically, all TGF-β superfamily sequences, except the TGF-β family and Müllerian inhibitory substance, are best aligned with TGF-β2 by using an alignment gap at the same finger 1 position as

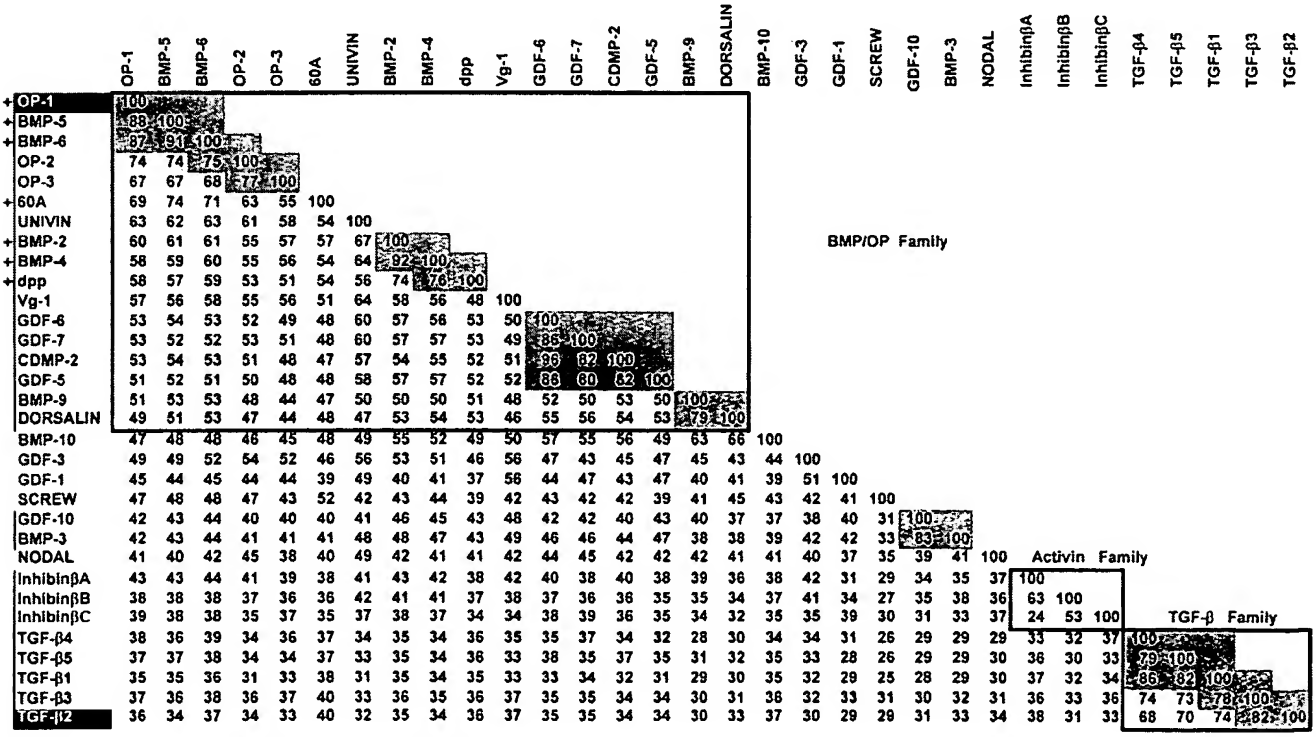
OP-1. This suggests that the conformation of the OP-1 Ω loop is more characteristic of the superfamily as a whole.

Based on the structure-based sequence alignment described above, we have constructed a matrix of the resulting percent identity between family member sequences (Fig. 6). In this set of proteins, OP-1 and TGF-β2 have the lowest sequence identity, only 36%, yet they are structurally very similar, suggesting that these molecules have evolved from a common ancestor. Secondary structure analyses of the sequences of OP-1 and TGF-β2 show that the heel region of OP-1 has a propensity to form an α-helix, whereas in TGF-β2 it does not. Stability of the helical conformation of the TGF-β2 heel region, therefore, requires interaction of this helix with the finger regions. As noted in Fig. 3, the specific residue positions involved in this interaction are essentially the same for both of these structures. Indeed, residues at these positions vary little in their physical and chemical properties for all the proteins listed in Fig. 6. We conclude that all of the TGF-β superfamily proteins in Fig. 6 share the OP-1/TGF-β2 structural motif.

Implications for Receptor Binding Specificity. The members of the TGF-β superfamily exert their biological effects by inducing the association of specific type I and type II serine/threonine kinase receptors. Various members have been shown to bind specifically to type II receptors, which then complex with type I receptors to initiate signal transduction (40).

It has been suggested that the receptor binding regions of TGF-β2 include residues that are both solvent-accessible and lie at positions of heterogeneous composition in aligned superfamily sequences (9). Divergent structural features in both OP-1 and TGF-β2 occur primarily in the external loops of finger 1 and finger 2, the loops bordering the helix, and the residues in the N-terminal domain preceding the first Cys of the cystine knot. These regions are also solvent-accessible. In both the OP-1 and TGF-β2 dimer structures, the tip of finger 2 and the Ω loop of finger 1 from one chain, and the C-terminal end of the α-helix in the heel of the other chain, form a contiguous ridge approximately 40 Å long and 15 Å wide (Fig. 5*A*). We propose that this ridge contains the primary structural features that interact with the receptor and that binding specificity between different TGF-β superfamily members derives from conformational and electrostatic variations on the surface of this ridge. Differences in the conformation of the finger 1 Ω loop, which constitutes the midsection of the ridge, and in the turn at the end of finger 2, which forms one end of the ridge, were noted. More significantly, there are striking differences in the surface charge of this region in OP-1 relative to TGF-β2 (Fig. 5 *B* and *C*). The ends of the finger regions are negatively charged in OP-1 whereas they are positively charged in TGF-β2. This results in a net charge of −4 for the receptor binding ridge of OP-1 versus +3 for TGF-β2. Conversely, the β-strand located C-terminal to the turn of finger 2 (β7, Fig. 1) is positively charged in OP-1, whereas it is negatively charged in TGF-β2 (Fig. 5 *B* and *C*). These observations suggest that electrostatic charge plays an important role in the specific interactions of OP-1 and TGF-β2 with their receptors and thus provide a molecular rationale for the functional differences between these two proteins.

We are grateful to Prof. Gregory Petsko for his enthusiastic support of this project, insightful discussions, and providing unlimited use of the crystallographic facilities at Brandeis University. We are also



	OP-1	BMP-5	BMP-6	OP-2	OP-3	60A	UNIVIN	BMP-2	BMP-4	dpp	Vg-1	GDF-6	GDF-7	CDMP-2	GDF-5	BMP-9	DORSALIN	BMP-10	GDF-3	GDF-1	SCREW	GDF-10	BMP-3	NODAL	InhibinβA	InhibinβB	InhibinβC	TGF-β4	TGF-β5	TGF-β1	TGF-β3	TGF-β2
+ OP-1	100																															
+ BMP-5	88	100																														
+ BMP-6	87	91	100																													
OP-2	74	74	75	100																												
OP-3	67	67	68	77	100																											
+ 60A	69	74	71	63	55	100																										
UNIVIN	63	62	63	61	58	54	100																									
+ BMP-2	60	61	61	55	57	57	67	100																								
+ BMP-4	58	59	60	55	56	54	64	92	100																							
+ dpp	58	57	59	53	51	54	56	74	76	100																						
Vg-1	57	56	58	55	56	51	64	58	56	48	100																					
GDF-6	53	54	53	52	49	48	60	57	56	53	50	100																				
GDF-7	53	52	52	53	51	48	60	57	57	53	49	86	100																			
CDMP-2	53	54	53	51	48	47	57	54	55	52	51	96	82	100																		
GDF-5	51	52	51	50	48	48	58	57	57	52	52	86	80	82	100																	
BMP-9	51	53	53	48	44	47	50	50	50	51	48	52	50	53	50	100																
DORSALIN	49	51	53	47	44	48	47	53	54	53	46	55	56	54	53	79	100															
BMP-10	47	48	48	46	45	48	49	55	52	49	50	57	55	56	49	63	66	100														
GDF-3	49	49	52	54	52	46	56	53	51	46	56	47	43	45	47	45	43	44	100													
GDF-1	45	44	45	44	44	39	49	40	41	37	56	44	47	43	47	40	41	39	51	100												
SCREW	47	48	48	47	43	52	42	43	44	39	42	43	42	42	39	41	45	43	42	41	100											
GDF-10	42	43	44	40	40	40	41	46	45	43	48	42	42	40	43	40	37	37	38	40	31	100										
BMP-3	42	43	44	41	41	41	48	48	47	43	49	46	46	44	47	38	38	39	42	42	33	83	100									
NODAL	41	40	42	45	38	40	49	42	41	41	42	44	45	42	42	42	41	41	40	37	35	39	41	100								
InhibinβA	43	43	44	41	39	38	41	43	42	38	42	40	38	40	38	39	36	38	42	31	29	34	35	37	100							
InhibinβB	38	38	38	37	36	36	42	41	41	37	38	37	36	36	35	35	34	37	41	34	27	35	38	36	63	100						
InhibinβC	39	38	38	35	37	35	37	38	37	34	34	38	39	36	35	34	32	35	35	39	30	31	33	37	24	53	100					
TGF-β4	38	36	39	34	36	37	34	35	34	36	35	35	37	34	32	28	30	34	34	31	26	29	29	29	33	32	37	100				
TGF-β5	37	37	38	34	34	37	33	35	34	36	33	38	35	37	35	31	32	35	33	28	26	29	29	30	36	30	33	79	100			
TGF-β1	35	35	36	31	33	38	31	35	34	35	33	33	34	32	31	29	30	35	32	29	25	28	29	30	37	32	34	86	82	100		
TGF-β3	37	36	38	36	37	40	33	36	35	36	37	35	35	34	34	30	31	36	32	33	31	30	32	31	36	33	36	74	73	78	100	
TGF-β2	36	34	37	34	33	40	32	35	34	36	37	35	35	34	34	30	33	37	30	29	29	31	33	34	38	31	33	68	70	74	82	100

FIG. 6. Identity matrix for the TGF-β superfamily. Included in the matrix are members that are more than 36% identical to OP-1 in their 7-Cys domains. Sequences were positioned in the matrix with those more distant from OP-1 at the bottom and arranged so as to cluster those of high identity. Boxes enclose families of sequences having 50% or higher identity with a majority of the other members of the family; gray areas identify groups of sequences having 75% or higher identity. Recombinantly expressed OP/BMP family members known to be able to make bone are denoted by a + symbol to the left. Superfamily members whose three-dimensional structures have been determined (OP-1 and TGF-β2) are highlighted black to the left. Some superfamily members were not included in the matrix because their level of sequence identity with either OP-1 or TGF-β2 was too low to permit meaningful structure-inferred sequence alignment (i.e., Müllerian inhibitory substance, MIS). The sequences are referenced in Kingsley (7), except as follows: UNIVIN (31), SCREW (32), BMP-9 (33), BMP-10 (34), GDF-5 (35) [also called CDMP-1 (36)], GDF-6 (35), GDF-7 (35), CDMP-2 (36), OP-3 (37), inhibin βC (38), and GDF-10 (39). Several sequences in the matrix have alternate names: OP-1(BMP-7), BMP-2(BMP-2a), BMP-4(BMP-2b), BMP-6(Vgr1), OP-2(BMP-8), 60A(Vgr-D), BMP-3 (osteogenin), GDF-6 (BMP-13), and GDF-7 (BMP-12).

grateful to James Griffith for expert advice and help. We also acknowledge Dan Peisach for generating the images in Figs. 1 and 2, Lynn Rardin for help with data collection, Jim Huston and John Smart for enthusiasm and support, and Charlie Cohen for critical review of the manuscript. This work was supported by a grant from Creative BioMolecules.

1. Sampath, T. K., Coughlin, J. E., Whetstone, R. M., Banach, D., Corbett, C., Ridge, R. J., Özkaynak, E., Oppermann, H. & Rueger, D. C. (1990) *J. Biol. Chem.* **265,** 13198–13205.
2. Cook, S. D., Wolfe, M. W., Salkeld, S. L. & Rueger, D. C. (1995) *J. Bone Joint Surg.* **77A,** 734–750.
3. Özkaynak, E., Schnegelsberg, P. N. J. & Oppermann, H. (1991) *Biochem. Biophys. Res. Commun.* **179,** 116–123.
4. Helder, M. N., Özkaynak, E., Sampath, K. T., Luyten, F. P., Latin, V., Oppermann, H. & Vukicevic, S. (1995) *J. Histochem. Cytochem.* **43,** 1035–1044.
5. Perides, G., Jensen, F. E., Edgecomb, P., Rueger, D. C. & Charness, M. E. (1995) *Neurosci. Lett.* **187,** 21–24.
6. Lefer, A. M., Psau, P. S., Ma, X.-L. & Sampath, T. K. (1992) *J. Mol. Cell. Cardiol.* **24,** 585–593.
7. Kingsley, D. (1994) *Genes Dev.* **8,** 133–146.
8. Daopin, S., Piez, K. A., Ogawa, Y. & Davies, D. R. (1992) *Science* **257,** 369–373.
9. Schlunegger, M. & Grütter, M. (1992) *Nature (London)* **358,** 430–434.
10. Özkaynak, E., Rueger, D. C., Drier, E. A., Corbett, C., Ridge, R. J., Sampath, K. T. & Oppermann, H. (1990) *EMBO J.* **9,** 2085–2093.
11. Sampath, K. T., Maliakal, J. C., Hsuschka, T., Jones, W. K., Sask, H., Tucker, R. F., White, K. H., Coughlin, J. E., Tucker, M. M., Pang, R. H. L., Corbett, C., Özkaynak, E., Oppermann, H. & Rueger, D. C. (1992) *J. Biol. Chem.* **267,** 20352–20362.
12. Griffith, D. L., Oppermann, H., Rueger, D. C., Sampath, T. K., Tucker, R. F. & Carlson, W. D. (1994) *J. Mol. Biol.* **244,** 657–658.
13. Higashi, T. (1990) R-AXIS-IIC, A Program for Indexing and Processing R-AXIS IIC Imaging Plate Data (Rigaku, Danvers, MA).
14. Collaborative Computation Project (1994) *Acta Crystallogr. D* **50,** 760–763.
15. Ten Eyck, L. F. & Arnone, A. (1976) *J. Mol. Biol.* **100,** 3–11.
16. Wang, B. C. (1985) *Methods Enzymol.* **115,** 90–112.
17. Reed, R. J. (1986) *Acta Crystallogr. A* **42,** 140–149.
18. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, N. (1991) *Acta Crystallogr. A* **47,** 110–119.
19. Brunger, A. T., Kuriyan, J. & Karplus, M. (1987) *Science* **235,** 458–460.
20. Lee, B. & Richards, F. M. (1971) *J. Mol. Biol.* **55,** 379–400.
21. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) *J. Mol. Biol.* **112,** 535–542.
22. Gilson, M. K. & Honig, B. H. (1987) *Nature (London)* **330,** 84–86.
23. Nicholls, A. & Honig, B. (1991) *J. Comput. Chem.* **12,** 435–445.
24. McDonald, N. Q. & Hendrickson, W. A. (1993) *Cell* **73,** 421–424.
25. Oefner, C., D'Arcy, A., Winkler, F. K., Eggimann, B. & Hosang, M. (1992) *EMBO J.* **11,** 3921–3926.
26. Lapthorn, A. J., Harris, D. C., Litlejohn, A., Lustbader, J. W., Canfield, R. E., Machin, K. J., Morgan, F. J. & Isaacs, N. W. (1994) *Nature (London)* **369,** 455–461.
27. Wu, H., Lustbader, J. W., Liu, Y., Canfield, R. E. & Hendrickson, W. A. (1994) *Structure Curr. Biol.* **2,** 545–558.
28. Leszczynski, J. F. & Rose, G. D. (1986) *Science* **234,** 849–855.
29. Sibanda, B. L. & Thornton, J. M. (1991) *Methods Enzymol.* **202,** 59–82.
30. McDonald, N. Q., Lapatto, R., Murray-Rust, J., Gunning, J., Wlodawer, A. & Blundell, T. L. (1991) *Nature (London)* **354,** 411–414.
31. Stenzel, P., Angerer, L. M., Smith, B. J., Angerer, R. C. & Vale, W. W. (1994) *Dev. Biol.* **166,** 149–158.
32. Arora, K., Levine, M. S. & O'Connor, M. B. (1994) *Genes Dev.* **8,** 2588–2601.
33. Wozney, J. M. & Celeste, A. (1993) Eur. Patent Appl. WO 93/00432, seq. id. no. 9.
34. Celeste, A. & Wozney, J. M. (1994) Eur. Patent Appl. WO 94/26893, seq. id. no. 1.
35. Storm, E. E., Huynh, T. V., Copeland, N. G., Jenkins, N. A., Kingsley, D. M. & Lee, S. J. (1994) *Nature (London)* **368,** 639–643.
36. Chang, S. C., Hoang, B., Thomas, J. T., Vukicevic, S., Luyten, F. P., Ryba, N. J. P., Kozak, C. A., Reddi, A. R. & Moos, M. (1994) *J. Biol. Chem.* **269,** 28227–28234.
37. Özkaynak, E. & Oppermann, H. (1994) Eur. Patent Appl. WO 94/10203, seq. id. no. 1.
38. Hötten, G., Neifhardt, H., Schneider, C. & Pohl, J. (1995) *Biochem. Biophys. Res. Commun.* **206,** 608–613.
39. Cunningham, N. S., Jenkins, N. A., Gilbert, D. J., Copeland, N. G., Reddi, A. H. & Lee, S.-J. (1995) *Growth Factors* **12,** 99–109.
40. Massague, J., Attisano, L. & Wrana, J. L. (1994) *Trends Cell Biol.* **4,** 172–178.